DOCUMENT RESUME

ED 445 024 TM 031 629

AUTHOR Williamson, David M.; Bejar, Isaac I.
TITLE Kohonen Self-Organizing Maps in Validity Maintenance for Automated Scoring of Constructed Response.
PUB DATE 2000-04-00
NOTE 26p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 25-27, 2000)
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classification; *Constructed Response; *Grading; *Scoring; Test Scoring Machines; *Validity
IDENTIFIERS Monitoring

ABSTRACT

As the automated scoring of constructed responses reaches operational status, monitoring the scoring process becomes a primary concern, particularly if automated scoring is intended to operate completely unassisted by humans. Using actual candidate selections from the Architectural Registration Examination (n=326), this study uses Kohonen Self-Organizing Maps (SOM) to build on previous research (D. Williamson, A. Hone, S. Miller, and I. Bejar, 1998) suggesting that classification trees are a useful means of validity maintenance. Classification trees can assist in identifying sources of disagreement between human and automated scoring, identify tendencies for human graders to overlook elementary or complex solutions, and provide significant efficiency in future case selection for human intervention. Since classification trees require a criterion value of score discrepancy between human and automated scores, Kohonen SOM provide an advantage in the ability to classify solutions in similar groups through neural networks without requiring prior human grading. Results suggest that Kohonen SOM could be used to classify solutions prior to human grading and classification tree analyses, thus providing a 43% reduction in the human grading required. However, further analyses are needed to establish whether classification trees would produce similar results with a reduced sample on the basis of Kohonen SOM classifications. (Contains 3 figures, 3 tables, and 14 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Running head: KOHONEN SELF-ORGANIZING MAPS

ED 445 024

Kohonen Self-Organizing Maps in Validity Maintenance for Automated Scoring of Constructed Response

David M. Williamson

The Chauncey Group International

Isaac I. Bejar

Educational Testing Service

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

D.M. Williamson

1

TM031629

Presented at the annual meeting of the National Council on Measurement in Education, New Orleans, Louisiana, April, 2000.

BEST COPY AVAILABLE

Abstract

As the automated scoring of constructed responses reaches operational status (e.g. Kenney, 1997, Burnstein, Kukich, Wolff, & Lu, 1998) monitoring the scoring process becomes a primary concern, particularly if automated scoring is intended to operate completely unassisted by humans. Using actual candidate solutions from the Architectural Registration Examination (ARE) this study utilizes Kohonen Self-Organizing Maps (SOM) to build on previous research (Williamson, Hone, Miller, & Bejar, 1998) suggesting that classification trees (Breiman, Friedman, Oshen, & Stone, 1984) are a useful means of validity maintenance. Specifically, that classification trees can assist in identifying sources of disagreement between human and automated scoring, identify tendencies for human graders to overlook elements of complex solutions, and to provide significant efficiencies in future case selection for human intervention. However, since classification trees require a criterion value of score discrepancy between human and automated scores Kohonen SOM provide an advantage of ability to classify solutions in similar groups through neural networks without requiring prior human grading. The results suggest that the use of Kohonen SOM could be used to classify solutions prior to human grading and classification tree analyses, thus providing a 43% reduction in the human grading required. However, further analyses are needed to establish whether classification trees would produce similar results with a reduced sample on the basis of Kohonen SOM classifications.

Kohonen Self-Organizing Maps in Validity Maintenance for Automated Scoring of Constructed Response

As automated scoring of complex constructed responses becomes operational for high-stakes assessments (e.g., Kenney, 1997 in architecture, Burnstein, Kukich, Wolff, & Lu, 1998 in text essays, Clauser, Margolis, Clyman, & Ross, 1997 in medicine) validation efforts, like the items themselves, become more complex than for traditional assessment methodologies (see Williamson, Bejar, & Hone, 1999). Typically, validation of automated scoring for complex constructed-response items includes a process of detailed examination of resultant scores in comparison with the scores of expert human graders (e.g. Clauser, Margolis, Clyman, & Ross, 1997; Sebrechts, Bennett, & Rock, 1991; Williamson, Bejar, & Hone, 1999). Since the development of automated scoring methodologies are typically conducted using a limited set of developmental data ongoing efforts to ensure that automated scoring continues to perform in an appropriate fashion is critical to the continued success of automated scoring for the full range of actual candidate solutions. The term *validity maintenance* is used to describe ongoing efforts to ensure the continued validity of an established operational program. This includes such efforts as:

- The identification and modification of aspects of automated scoring not performing as intended
- The identification of areas where automated scoring may be performing as intended but may benefit from further review with subsequent

BEST COPY AVAILABLE

determination of whether changes should be incorporated in future releases and implementation of such changes

- Critical review of the adequacy of tutorial software, implications for examination task performance, and modifications recommended for improving tutorials
- Evaluation and modification of instructions, interface operation, diagrams and multimedia applications within the examination and implications for resultant scores
- The identification of scoring criteria that may be appropriate for future inclusion/exclusion from automated scoring
- The identification of new processes and procedures, both developmental and analytical, which may be applied to ongoing evaluation of the examination.

The term *first-order* validity maintenance refers to investigations of the immediate integrity of automated scoring: is it performing in the manner in which it was intended? The potential impact of any malfunction on resultant candidate scores may require human intervention to correct for malfunctioning aspects, either in individual features or overall scores, of automated scoring. Clearly, when some aspect of automated scoring is malfunctioning it should be fixed as soon as possible. In practice, a variety of circumstances (e.g. programming limitations, interface limitations, software packaging and distribution, etc.) may make it difficult to immediately institute a correction to such malfunctioning aspects of automated scoring. Therefore, there is significant value not only in identifying any malfunctioning scoring features but also in efficiently identifying

BEST COPY AVAILABLE

solutions that may be affected by the malfunction and implementing corrective procedures.

In contrast, the term *second-order* validity maintenance addresses the long-term precision and evolution of automated scoring of complex constructed responses. Issues identified in second-order validity maintenance procedures are those in which automated scoring is performing as it was intended to perform but a committee of experts reviewing the solutions may suggest some scoring adjustments to better reflect their opinions on particular issues. That is, while first-order validity maintenance focuses on issues of quality assurance, second-order validity maintenance is concerned with the identification of potential validity enhancements from changes to automated scoring. Examples of these types of issues may include different recommended weightings of criteria, different tolerance for less-than-perfect implementations of criteria, and inclusion or exclusion of criteria that may be marginally or tangentially related to the purpose of the particular constructed-response task. Of course, any two groups of experts will disagree on some aspects of practice so the findings from second-order validity maintenance processes can best be considered as 'suggestions' rather than as 'problems' with automated scoring, which would be the domain of first-order validity maintenance. The process of second-order validity maintenance can help assure that all reasonable solution implementations are evaluated appropriately by the automated scoring as well as providing possibilities for the future evolution of the examination and the automated scoring.

Such ongoing endeavors of validity maintenance can require a substantial degree of time and effort. Therefore, these processes would benefit from procedures that may focus attention on the most fruitful sources of potential scoring process adjustments. A

BEST COPY AVAILABLE

previous study (Williamson, Hone, Miller, & Bejar, 1998) investigated the use of Classification and Regression Trees (CART) for applications in validity maintenance for a vignette from the Architect Registration Examination (ARE). In that study, the goal of the analysis was to identify the pattern of features associated with different degrees of agreement between human and computer scoring. Specifically, the difference between human and computer scores were regressed against all the features extracted as part of the automated scoring process. The Williamson et al (1998) study found that CART techniques can be a promising method of identifying issues of validity maintenance as well as being useful for the automated case selection of future cases which require first-order validity maintenance interventions. Because CART is a regression/classification method it is necessary to have an existing set of solutions with both automated and human scores to serve as the criterion for training the CART trees for the identification of cases suspected of possessing validity maintenance issues.

CART is a *supervised learning* method. That is, once the learning has taken place and the regression tree has been estimated and validated, new cases for which no human scores are available can be classified into their likely level of disagreement with human scoring. By contrast Kohonen Self-Organizing Maps (SOM) are an instance of *unsupervised* learning. That is, whereas the CART method uses the extracted features and a disagreement variable in the analysis SOMs are estimated from the features only. Therefore, SOMs may provide an advantage in efficiency by substantially reducing the need for human graders. This study evaluates the potential for Kohonen SOM to reduce the need for human graders in the validity maintenance process by comparing the

BEST COPY AVAILABLE

application of Kohonen SOM without a criterion value to the previously obtained results of the CART techniques.

Overview of the ARE

The ARE consists of nine divisions, six of which are multiple-choice examinations and three of which are fully computerized graphic simulations of architectural design tasks. The candidate receives either a "Pass" or "Fail" for each division and must pass all divisions to become registered. Each of the three graphic divisions is made up of a series of small, focused design problems called "vignettes". There are a total of 15 different vignettes that comprise the three graphic divisions of the ARE. (More information on the administration and scoring of the ARE may be found in Bejar & Braun, 1994; 1999; and Kenney, 1997).

The scoring of the vignettes comprising the three graphic divisions of the ARE, referred to as a 'mental model' approach (see Williamson, Bejar & Hone, 1999), are the culmination of studies of the criteria, judgments, weighting, and considerations of expert human graders during holistic grading processes for operational paper-and-pencil vignettes (Bejar, 1991; Oltman, Bejar, & Kim, 1993; Bejar & Braun, 1994). This 'mental model' scoring incorporates several aspects of human scoring processes, the most obvious of these aspects being the division of each candidate's vignette solution into specific and distinct elements of scoring criteria, 'features', with each feature receiving an evaluation of A (acceptable), I (indeterminate), or U (unacceptable). The I represents a borderline implementation. These features are aggregated to a summary evaluation of A, I or U on the vignette that mimics the aggregation processes utilized by expert human graders.

Overview of previous CART study

A previous study (Williamson, Hone, Miller, & Bejar, 1998) utilized CART in validity maintenance to evaluate the utility of classification trees (Breiman, Friedman, Oshen, Stone, 1984) for performing first-order and second-order validity maintenance processes. A specific goal was to automate the identification of cases where experienced graders and automated scoring can be expected to disagree as a result of automated scoring malfunction (first-order validity maintenance).

As in linear regression and discriminant function analyses, the analysis required data (often called "training" data) on the attributes (or independent variables, in this case the solution 'features') and the classification outcome (or dependent variable, in this case the human-automated score agreement). Unlike linear regression analysis, where the outcome is a prediction equation, the outcome of CART is a binary tree. A binary tree consists of a set of sequential binary decisions, applied to each case, that lead to further binary decisions or to a final classification (terminal node) of a that case. A resultant tree can be used to classify new cases where the dependent variable is not available. Given a classification tree, new cases are "filtered down" the tree to a final classification. A classification tree based on a classic data set (Iris flower species) used by R.A. Fisher to illustrate discriminant analysis is provided as Figure 1. This classification tree utilizes critical values of petal width and petal length to classify the cases into one of three terminal nodes, each representing a different species of Iris. These terminal classification nodes may be characterized in table format by decision vectors that represent the decision sequence and outcome of the classification tree. The decision vectors corresponding to the Iris classification tree in Figure 1 are presented as Table 1.

For the analyses in the Williamson et al (1998) study the CART methods generally proved to be fruitful approaches to both first-order and second-order validity maintenance for the initial 'training' set of 326 actual candidate solutions for a vignette from the ARE for which both human and automated scores were available. A difference score was computed by subtracting the numeric value of the automated score from the numeric value of the human holistic score and was used as the criterion value for CART. The possible values are provided as Table 2.

In applications directed at first-order validity maintenance these methods indicated specific features which required intervention and suggested others which upon investigation provided evidence about the advantage of specificity and thoroughness provided by automated scoring systems. Examinations with respect to second-order validity maintenance processes revealed aspects which may be worthy of consideration for the continued evolution of automated scoring as well as giving some indication of the frequency and conditions for which these possibilities may be relevant.

The feature vectors for the CART analyses which represent the classification tree produced are provided as Table 3 (in much the same way that Table 1 represents Figure 1). In the architectural evaluation of solutions with these feature vectors the following conclusions were drawn about selected feature vectors:

- feature vector A – revealed two minor features considered by the human graders and not by automated scoring and slight differences in feature weighting
- feature vector B – revealed a difference in feature weighting between human and automated scoring

BEST COPY AVAILABLE

- feature vector C – revealed a difference in feature weighting between human and automated scoring
- feature vector M – human graders were overlooking an important feature
- feature vector N – human graders were making allowances for candidate misinterpretation

Classification trees were produced for the purpose of selecting future cases requiring first-order validity maintenance interventions (potential score adjustments) from an additional set of 1117 actual candidate solutions which did not have human-produced grades. The feature vector from this use of classification trees was effective in the selection of all future cases requiring first-order validity maintenance intervention while reducing the burden of the review process by 68% over the policy of reviewing 100% of the solutions for cases requiring intervention.

Overview of Kohonen Self-Organizing Maps

Kohonen SOM are a neural network methodology closely related to cluster analysis and other dimensionality reduction methods in that it is a methodology for unsupervised learning, that is, learning without reference to a criterion value or prior information. As such it functions to build representations of complex data by representing the data as a one- or two-dimensional array of artificial neurons.

As a neural network technique it mimics the manner in which certain tasks are represented in the brain as spatially ordered groups of neurons which respond to certain types of stimulus. Therefore, Kohonen SOM make the assumption that clusters (or classes) are formed from patterns that share common features. Each artificial neuron is a representation of configurations of patterns in data.

BEST COPY AVAILABLE

Initially these artificial neurons are set to starting values. In processing complex data, such as data from complex constructed response tasks, each case becomes associated with the particular neuron in the grid of artificial neurons that is most like the case in question. The artificial neuron in question then has its weights altered to reflect the contribution of the new input pattern from the data. In an iterative process of reconsidering the input data and readjusting the network of artificial neurons to reflect the new associations of input data to the artificial neurons different areas of the Kohonen SOM respond to different types of patterns within the data.

Ultimately, a grid of neurons is produced with each neuron associated more or less strongly with particular input data pattern. This grid of neurons and the cases comprise a pattern which might be conceptualized as a contour map of the data, with groups of similar cases clustering together in specific neurons or SOM regions and the number of cases associated with a particular neuron determining the 'height' at various SOM points.

For continuity with the CART example above utilizing Fisher's Iris data (Figure 1 and Table 1) an example of the output classification from a 1x3 Kohonen SOM for the Iris data is provided as Figure 2. In Figure 2 each box represents the location of a particular numbered neuron in the 1x3 Kohonen SOM. Immediately beneath the neuron number is the number of Iris flowers which are associated with that neuron. Beneath these neuron sample sizes are the Iris species (Setosa, Versicolor, and Virginica) corresponding to the neuron. In parentheses following each Iris species is the percentage of flowers of that species in the neuron.

BEST COPY AVAILABLE

To the extent that various regions of the SOM correspond to certain aspects of interest in the data such techniques permit rapid and efficient organization of complex data for a variety of purposes. In addition, the representation of the data in a SOM provides a means for swiftly visualizing trends and tendencies within complex data. (For more about Kohonen SOM see Balakrishan Cooper, Jacob & Lewis. 1994; Murtagh & Hernandez-Pajares, 1995; or Waller, Kaiser, Illian, & Manry, 1998).

Results and Discussion

The data for this study consisted of the same 326 actual candidate solutions for a particular ARE vignette which were used as the 'training' set in the original CART study. These data were analyzed utilizing Neural Connection 2.0 (SPSS Inc./Recognition Systems Inc.). Only 261 (80%) of the 326 candidate solutions were utilized in developing the Kohonen SOM in order to preserve 20% of the sample for additional validation and testing analyses. The neural network was specified to consist of a 5x5 neuron grid for a total of 25 possible neurons in the net. The composition of the resultant neuron clusters were compared to the previous CART feature vector results from the Williamson et al study.

The correspondence between the results of the Kohonen SOM analysis and the original CART feature vectors (from table 3) are provided as Figure 3. In Figure 3 each box represents the location of a particular numbered neuron in the 5x5 Kohonen SOM. Immediately beneath the neuron number is the number of solutions which are associated with that neuron. Beneath these neuron sample sizes are the feature vector codes (letters A through N from Table 3) from the CART analysis corresponding to the neuron. In

BEST

parentheses following each feature vector code is the percentage of solutions in that neuron which have that feature vector.

It is initially striking that a single neuron (neuron 25) contains 45% of the total sample of cases while the remaining neurons contain no more than 7% of the sample cases. However, the fact that all of the cases in neuron 25 correspond to CART feature vector J, in which the human graders and the automated scoring were in perfect agreement, provides some reassurance that this peak in the neural net is an appropriate representation of the data. It is certainly not surprising to find that there exists a dramatic peak corresponding to feature vectors indicative of perfect agreement between automated and human scoring when the validity of the automated scoring has already been established through such comparisons (e.g. Williamson, Bejar, & Hone, 1999). With the contribution of neuron 25 the lower right-hand region of the Kohonen SOM (neurons 20, 23 and 25) represents a region of agreement between human and automated scores. This region comprises 97% of all feature vector J's in the sample and 84% of all feature vectors representing complete agreement in the sample.

Neuron 5 represents the cases that correspond to feature vector M from the CART analyses, with 92% of all feature vector M cases associated with this neuron. Similarly, neurons 3 and 8 are associated with feature vector N from the CART analyses, with 81% of all feature vector N cases associated with these two neurons. Taken together (and ignoring the three cases which appear in neuron 4) these suggest that the upper right-hand region of the Kohonen SOM (neurons 3, 5 and 8) represent a region in which the human graders were somewhat more lenient than the automated scoring (ie automated score of I and human score of A or automated score of U and human score of I). Yet, even with

BEST COPY AVAILABLE

this region of the network associated with this interpretation the separation between the M and N vectors, and their separate interpretations (from the CART study), remain distinct.

Neurons 13 and 18 are associated with feature vector A from the CART analyses, with 85% of all cases with this feature vector associated with these two neurons. Similarly, neuron 7 is associated with feature vector C from the CART analyses (with 58% of all cases of feature vector C) and neuron 1 is associated with feature vector B from the CART analyses (with 100% of all cases of feature vector B). Together these represent a diagonal region in the network which represents instances in which the human graders were more strict than the automated scoring (with human scores of U and automated scores of A). Yet, again within the larger region the former CART feature vectors remain in distinct pockets of the Kohonen SOM.

Neuron 15 is associated with feature vector F from the CART analyses, with 100% of all cases with this feature vector associated with this neuron. Since feature vector F was not emphasized in the determination of what the feature vectors represent in the CART analyses it is unclear what specific interpretation these solutions may have received, though they do represent cases in which the automated scoring was slightly more lenient than the human graders. This may be an important contribution of the Kohonen SOM to the validity maintenance process in that if this Kohonen SOM was produced prior to or concurrently with the CART analyses feature vector F may have received more attention in the subsequent solution reviews.

Finally, the lower left-hand region of the Kohonen SOM (neurons 11, 16, 17, 21 and 22) may be vaguely described as being a region in which human and automated

BEST COPY AVAILABLE

scores are in complete agreement. Within this region 70% of the cases are consistent with feature vectors from the CART analyses indicating complete agreement. However, the number of cases associated with these neurons is sufficiently small to cast some doubt on this interpretation.

The results suggest that there is a reasonably high degree of concordance between the Kohonen SOM and the classification tree terminal nodes from CART analyses. The neurons (or pairs of neurons) in the Kohonen SOM tend to be associated with cases which are parallel to the cases classified by the individual feature vectors from CART. Furthermore, the regions of the Kohonen SOM tend to be consistent with particular classes of human and automated scoring agreement, with the lower right-hand region being a region of agreement, the upper right-hand region being cases where the human grading was more lenient than automated scoring, and the diagonal from the center to the top left-hand corner being a region of cases in which the automated scoring was more lenient than human grading.

Given this consistency between the neurons and regions of the Kohonen SOM (an unsupervised method) and the classification trees from CART analysis (requiring a criterion score), the use of Kohonen SOM could provide additional efficiencies in the validity maintenance process by prescreening data for sampling human scores. For example, for this sample of 261 solutions if the Kohonen SOM was produced prior to any human grading, and thus any CART analyses, the resultant neuron clusters of solutions could have been used to provide a stratified sampling of solutions to receive human grades. If, for example, 15 solutions from each neuron (obviously less for neurons with less than 15 solutions) were selected to receive human grades then 149 (57%) of the

BEST COPY AVAILABLE

solutions would be human graded rather than the entire set of 261 solutions. Such a savings of 43% of human grading effort is not inconsequential since each human grading requires the effort of a committee of three graders and some degree of time as they evaluate the solution and discuss their observations to come to a determination of final score. Once the human scores were obtained on this stratified sample of solutions they could then be subjected to CART analyses on the basis of the smaller stratified sample to (presumably) produce a classification tree and feature vectors similar to those previously obtained with the larger data set. These CART results could then be utilized to the advantages suggested by the Williamson et al. (1998) study. While being insufficient for complete reliance these results suggest that Kohonen SOM are worthy of further investigation regarding their ability to substantially reduce the burden of human scoring in validity maintenance processes.

References

Balakrishan, P. V., Cooper, M. C., Jacob, V. S., & Lewis, P.A. (1994). A study of the classification capabilities of neural networks using unsupervised learning: A comparison with K-means clustering. Psychometrika, 59, 509-525.

Burstein, J. Kukich, K., Wolfi, S., & Lu, C. (1998). Computer analysis of essay content for automated score prediction. Presented at the Annual Meeting of the National Council of Measurement in Education. San Diego.

Bejar I. I. (1991). A methodology for scoring open-ended architectural design problems. Journal of Applied Psychology, 76, (4), 522-532.

Bejar, I. I., & Braun, H. (1994). On the synergy between assessment and instruction: Early lessons from computer-based simulations. Machine-Mediated Learning, 4, 5-25.

Bejar, I. I and Braun, H. I. (March, 1999). Architectural simulations: From research to implementation: Final report to the National Council of Architectural Registration Boards (Research Memorandum RM-99-2). Princeton, NJ: Educational Testing Service.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. Belmont, California: Wadsworth.

Clauser, B. E., Margolis, M. J., Clyman, S. G., & Ross, L. P. (1997). Development of automated scoring algorithms for complex performance assessments: A Comparison of two approaches. Journal of Educational Measurement, 34, 141-161.

Kenney, J. F. (1997). New testing methodologies for the Architect Registration Examination. CLEAR Exam Review, 8, (2), 23-28.

Murtagh, F., & Hernandez-Pajares, M. (1995). The Kohonen self-organizing map method: An assessment. Journal of Classification, 12, 165-190.

Oltman, P. K., Bejar, I. I., & Kim, S. H. (1993). An approach to automated scoring of architectural designs. In U. Flemming & S. van Wyk (Eds.), CAAD Futures 93 (pp. 215-224). Pittsburgh, PA: Elsevier Science Publishers B. V.

Sebrechts, M. M., Bennett, R. E., & Rock, D. A. (1991). Agreement between expert system and human raters' scores on complex constructed-response quantitative items. Journal of Applied Psychology, 76, 856-862.

Waller, N. G., Kaiser, H. A., Illian, J. B., Mantry, M. (1998). A comparison of the classification capabilities of the 1-dimensional Kohonen neural network with two partitioning and three hierarchical cluster analysis algorithms. Psychometrika, 63, 5-22.

Williamson D. M., Bejar I. I., & Hone, A. S. (1999). 'Mental model' comparison of automated and human scoring. Journal of Educational Measurement, 36, (2), 158-184.

Williamson, D. M., Hone, A. S., Miller, S., & Bejar, I. I. (1998). Classification Trees for Quality Control Processes in Automated Constructed Response Scoring. Presented at the Annual Meeting of the National Council of Measurement in Education. San Diego.

BEST COPY AVAILABLE

Table 1

Decision Vectors Corresponding to the Iris Classification Tree

	Node 1	Node 2
Classification	PETALLEN	PETALWID
1	<=2.45	
2	>2.45	<=1.75
3	>2.45	>1.75

Table 2

Possible Difference Score Values (Human-Automated)

	Automated Score		
Human Score	A	I	U
A	0	1	2
I	-1	0	1
U	-2	-1	0

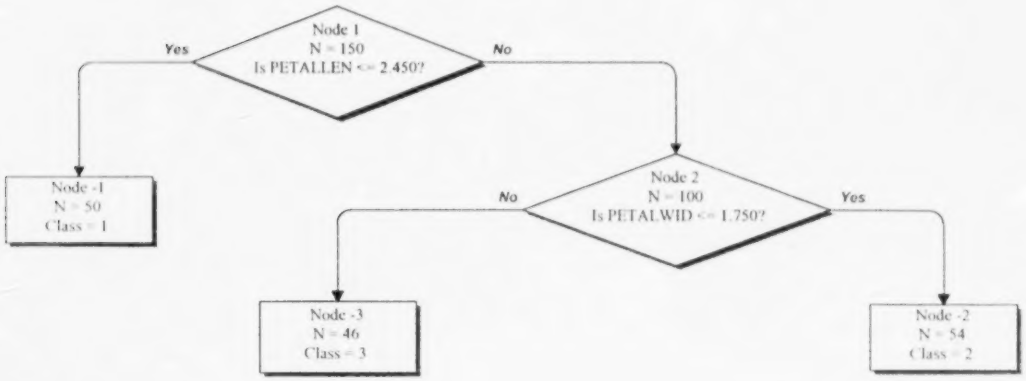
Table 3

Feature Vectors Utilizing the Difference Scores as the Dependent Variable

Terminal Node (Difference Score)	Feature Vector	N1 (F3)	N2 (F9)	N3 (F2)	N4 (F2)	N5 (F13)	N6 (F5)	N7 (F2)	N8 (F5)	N9 (F2)	N10 (F2)	N11 (F1)	N12 (F5)	N13 (F1)
-2	A	I,A	I,A	I										
-2	B	I,A	I,A	U,A	I,A	U								
-2	C	U								U,A	I,A	U,A		
-1	D	I,A	U					A	U					
-1	E	I,A	I,A	U,A	U									
-1	F	I,A	I,A	U,A	I,A	U								
-1	G	U								U,A	I,A	I		
-1	H	U								I			U	
0	I	I,A	U					U,I						
0	J	I,A	I,A	U,A	I,A	I,A	I,A							
0	K	U								U,A	U			
0	L	U								I			I,A	U,I
1	M	I,A	U					A	I,A					
1	N	U								I			I,A	A

Figure 1

Sample CART Analysis Classification Tree for the Iris Data



BEST COPY AVAILABLE

Figure 2

Kohonen SOM Result Classification of Fisher's Iris Data

Neuron 1	*Neuron 2*	*Neuron 3*
n=31	n=29	n=30
Versicolor (74%)	Versicolor (24%)	Setosa (100%)
Virginica (26%)	Virginica (76%)	

Figure 3

Kohonen SOM / CART Feature Vector Comparison

Neuron 1 n=12 B (50%) D (17%) H (33%)	*Neuron 2* n=0	*Neuron 3* n=10 C (30%) J (10%) N (60%)	*Neuron 4* n=3 K (100%)	*Neuron 5* n=19 C (11%) I (11%) K (11%) M (58%) N (11%)
Neuron 6 n=0	*Neuron 7* n=13 C (85%) G (15%)	*Neuron 8* n=13 K (8%) L (8%) N (85%)	*Neuron 9* n=0	*Neuron 10* n=1 C (100%)
Neuron 11 n=6 A (17%) J (50%) K (17%) M (17%)	*Neuron 12* n=0	*Neuron 13* n=19 A (74%) E (26%)	*Neuron 14* n=0	*Neuron 15* n=11 A (9%) F (45%) K (18%) L (9%) N (18%)
Neuron 16 n=4 I (25%) K (50%) L (25%)	*Neuron 17* n=2 C (50%) E (50%)	*Neuron 18* n=3 A (100%)	*Neuron 19* n=0	*Neuron 20* n=3 J (100%)
Neuron 21 n=5 A (20%) I (40%) K (40%)	*Neuron 22* n=3 E (33%) I (67%)	*Neuron 23* n=17 A (12%) C (6%) J (76%) N (6%)	*Neuron 24* n=0	*Neuron 25* n=117 J (100%)